E.N.S.S.I.B
ECOLE NATIONALE SUPERIEUR
DES SCIENCES DE L'INFORMATION
ET DES BIBLIOTHEQUES

DPSSIB

Diplôme Professionnel Supérieur en Sciences de l'Information et des Bibliothèques

Rapport de recherche bibliographique

LA REPRESENTATION DES CONNAISSANCES ET CONCEPTION D'UN SYSTEME D'INFORMATION

Présenté par:

NGABONZIMA Alphonse

Sous la direction de:

Mohamed HASSOUN
E.N.S.S.I.B

1995 - 1996



E.N.S.S.I.B

ECOLE NATIONALE SUPERIEUR
DES SCIENCES DE L'INFORMATION
ET DES BIBLIOTHEQUES

DPSSIB

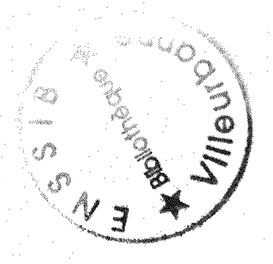
Diplôme Professionnel Supérieur en Sciences de l'Information et des Bibliothèques

Rapport de recherche bibliographique

LA REPRESENTATION DES CONNAISSANCES ET CONCEPTION D'UN SYSTEME D'INFORMATION

Présenté par:

NGABONZIMA Alphonse



Sous la direction de:

Mohamed HASSOUN
E.N.S.S.I.B

1995 - 1996

"La représentation des connaissances et conception d'un système d'information"

Alphonse NGABONZIMA

RESUME:

L'accès à l'information actualisée,pertinente et complète est devenu,de nos jours, un enjeu important pour les acteurs économiques, scientifiques et techniques.
Le meilleur moyen de répondre à ce besoin consiste à concevoir un système d'information performant,capable de répondre aux attentes des demandeurs de cette matière première.
La présente recherche bibliographique procède ainsi à répertorier la documentation sur la structuration,les méthodes et moyens de représenter les connaissances dans la conception d'un système d'information.

DESCRIPTEURS:

Représentation des connaissances, structures de données (informatique), base donnée, structuration, système information, conception système, modélisation, base connaissance, banque donnée.

ABSTRACT:

The accessibility of the informations update,important and complet is considered nowadays of grand importance for the economic,scientific and technical actors.A better way to answer the necessity consist of concept of a perfect information system which capable to answer the expectations of that primary material.
The aim of this bibliographical is to collect all documentation about structuration,methods and ways for to represent knowledge in concept of information system.

KEYWORDS:

Knowledge representation, datastructure,database, structuration, information system, system design, modeling, knowledge base, databank.

TABLE DE MATIERE

I - PRÉSENTATION DU SUJET

La représentation des connaissances et conception d'un système d'information est un sujet qui nous a été proposé par M. Mohamed HASSOUN, Maître de conférence à l'École Nationale des Sciences de l'Information et des Bibliothèques (ENSSIB).

De nos jours, d'aucuns si pas tous s'accordent que la capacité des sociétés à utiliser l'information est devenu le facteur le plus déterminant dans leur différence au niveau du dynamisme économique et intellectuel. La recherche étant le moteur du développement, et l'infirmation disponible l'outil essentiel à la recherche, c'est pourquoi bien des pays, diverses organisations et entreprises s'investissent dans l'organisation et la structuration de l'information en vue de la rendre plus accessible et rapidement exploitable.

Dans une brochure sur les banques de données produite par le Ministère de l'Enseignement Supérieur et de la Recherche, nous lisons :

> **Dans la conjoncture mondiale actuelle, fortement compétitive, l'accès à une information actualisée, pertinente et complète, est devenu un enjeu important pour les acteurs économiques, scientifiques et techniques. On assiste simultanément à une expansion considérable du marché de l'information et à une multiplication des supports et techniques de transmission de celle-ci.** *(1)*

Les systèmes d'information servant à collecter, traiter et distribuer les informations que les entreprises utilisent pour planifier, surveiller et contrôler leurs activités, à l'instar de l'évolution de ces organisations (entreprises ou organismes publics) le domaine de la représentation des connaissances et de la conception des systèmes d'information connaît lui aussi des innovations.

Face à cette réalité, notre sujet, **la représentation des connaissances et conception d'un système d'information** nous conduit à inventorier les documents pertinents qui traitent de ce thème surtout en ce qui est de la conception, la structuration, l'organisation et le traitement de l'information. Le but étant d'en proposer, le cas échéant, aux chercheurs qui sauraient utilement s'en servir.

---

(1) ***Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la recherche**, Des banques de données pour les étudiants, les enseignants, les chercheurs.- Paris 1991*

II - SYNTHÈSE

Selon l'encyclopaedia universalis, la notion de représentation est invoquée pour rendre compte du phénomène de la connaissance.

Connaître une chose, c'est se l'assimiler, se la rendre intérieure, tout en lui laissant son statut de réalité extérieure, indifférente, en tant que telle, au processus par lequel elle devient objet de connaissance *(2)*.

Représenter les connaissances, c'est ainsi pouvoir traduire les données numériques ou symboliques sous une forme utilisable par un système à base de connaissances ou mieux par un système d'information.

Les systèmes d'information, avions-nous déjà dit dans la présentation du sujet, servent à collecter, traiter et distribuer les informations que les entreprises utilisent pour planifier, surveiller et contrôler leurs activités.

Ceci étant, dans un système d'information, une bonne représentation du contenu d'un document doit tenir absolument compte des caractéristiques de son futur utilisateur.

D'où nécessité d'une conception conséquente d'un système d'information maîtrisable par les individus du système organisationnel. Le système d'information intervient d'une manière fondamentale dans le fonctionnement d'une organisation.

Forts de cette idée, beaucoup de chercheurs, notamment en France, se sont lancés dans la construction de théories avec l'objectif d'améliorer les méthodes de conception des systèmes d'information.

En 1984, dans la préface du livre de TARDIEU Hubert et ses collaborateurs - La conception d'un système d'information : construction de la base de données- Jean-Louis LEMOIGNE écrit : « Sur le thème de l'analyse et de la conception des systèmes d'information des organisations (entreprises ou administration), la littérature foisonne ». Mais il se plaint en disant que même les innovations les plus spectaculaires semblent n'affecter que très lentement les esprits et les pratiques en matière d'analyse et de conception de système *(3)*.

---

(2) Encyclopaedia Universalis, vol. 19 - Paris, 1990, pp 822-823

(3) Hubert TARDIEU et ali, Conception d'un système d'informations, les conditions d'organisation.- Paris, 1984, p 15

L'ambition traditionnelle était d'établir un schéma de référence suffisamment précis et suffisamment transparent pour que les spécialistes divers qui auraient à intervenir dans la conception et l'animation du système d'information disposent loyalement d'un même langage et d'un même mode d'emploi des mêmes outils, garant de la qualité et de l'efficacité des communications entre eux, et entre eux et le reste du monde. Ce schéma ne pouvait être légitime que dans le cas de processus industriels et robotisables. Il ne l'est pas, et ne peut pas l'être, dans les situations que créent les êtres humains en relation au sein de leurs organisations sociales.
Le système d'information n'est il pas aussi l'ensemble des moyens, humains et matériels, et des méthodes se rapportant au traitement des différentes formes d'informations rencontrées dans l'organisation.

Même si la tendance est d'éviter l'asservissement à un schéma rigoureux préétabli, en matière de système d'information il y a actuellement, pour les professionnels qui ont à les gérer, trois tâches essentielles : la conception, la réalisation et l'évolution.

Étant donné que nous avons des connaissances très limitées dans le domaine de la conception des systèmes d'information, nous essaierons de présenter synthétiquement ces tâches comme nous l'avons modestement compris après lecture des documents produits par des spécialistes en la matière *(4)* tels que :

- Hubert TARDIEU
- Jean-Louis CAVALERO
- Jean-Claude COURBON
- GALACSI

etc...

---

(4) - GALACSI, les systèmes d'information : analyse et conception. - Paris : Dunod, 1986, 280 p.

- Jean-Claude COURBON, systèmes d'information : structuration, modélisation et communication. - Paris : intéréditions, 1993, 288 p.

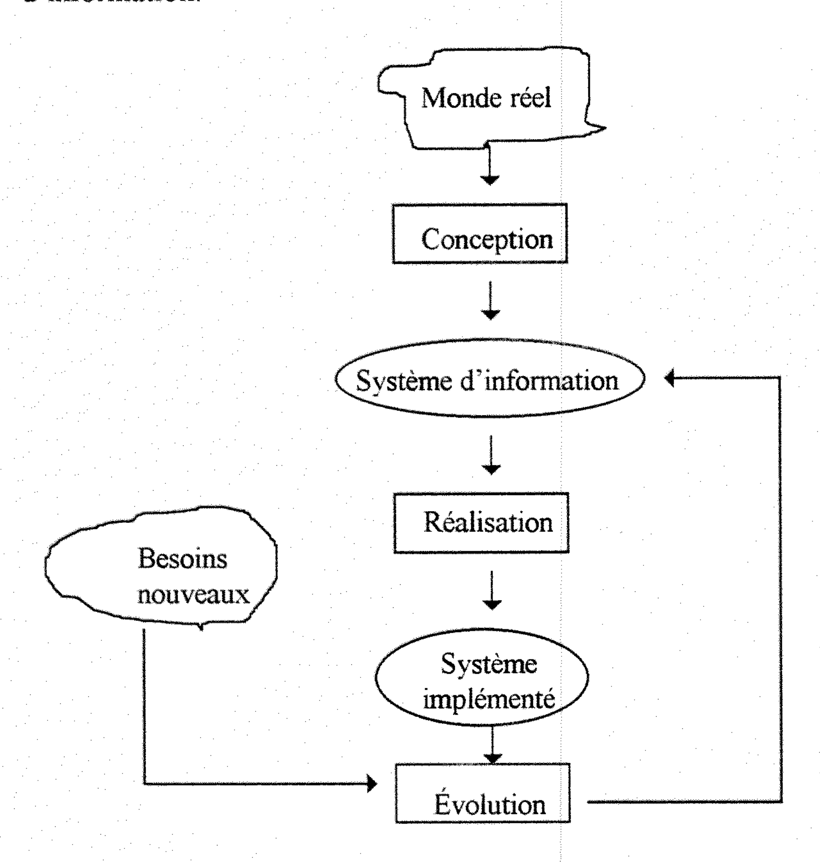
- Jean-Louis CAVARERO, La conception des systèmes d'information : un modèle, un dossier standard, des méthodes. - Paris : Masson, 1982, 188 p.

LA DÉMARCHE DE CONCEPTION D'UN SYSTÈME D'INFORMATION

Le travail du concepteur consiste à définir des objets et des règles de fonctionnement. Certains de ces objets et de ces règles se déduisent directement de l'analyse du monde réel : c'est la première étape de la conception, l'étape conceptuelle, qui correspond à la définition des invariants.

La seconde étape, l'étape logique consiste à définir les objets et les règles restants, c'est-à-dire les objets et les règles reflétant certaines vues, options ou contraintes de l'utilisateur.

Schéma sur : la conception, la réalisation et l'évolution d'un système d'information.



Que faut-il donner au concepteur pour lui permettre d'effectuer une bonne conception ?

La première chose à faire est de lui donner une liste des types d'objets et de règles qu'il aura à définir quel que soit le système d'information étudié.

Ces types d'objets et règles constituent les concepts des modèles de représentation du système d'information.

A - Le modèle de représentation des systèmes d'information

1. Le modèle de données

Grâce au modèle de données, le concepteur doit pouvoir définir à partir du monde réel la structure logique des informations manipulées dans le système d'information, c'est-à-dire la structure d'accès des fichiers ou des bases de données.
Trois concepts permettent de représenter le contenu du monde réel (niveau conceptuel) : l'entité, la rubrique, la relation.

Un concept permet de représenter la structure logique des informations : la structure d'accès (niveau logique).

♦ Le concept de rubrique

A chaque information du monde réel est associée une rubrique. Il existe deux catégories de rubriques : les rubriques élémentaires et les rubriques non élémentaires.

Une rubrique élémentaire représente le plus petit lot d'information manipulable dans le système d'information et ayant une signification intrinsèque.

Une rubrique non élémentaire (ou composée) est une rubrique obtenue par concaténation de rubriques élémentaires ou non.

Ex. : - la rubrique NOM - ABONNÉ est une rubrique élémentaire.
- la rubrique ADRESSE est une rubrique composée par la concaténation des deux rubriques :
. rue
. code postal

♦ Le concept d'entité

En première analyse, le monde réel n'est jamais examiné en terme de rubriques, mais en terme d'objets constitués par des ensembles de rubriques.

Une unité du monde réel est un objet, concret ou abstrait de la réalité caractérisé par des rubriques.

Ex : - L'entité ABONNÉ est caractérisée par les rubriques : RÉFÉRENCE-ABONNÉ, NOM-ABONNÉ, CODE POSTAL, etc...

- L'entité COMPTEUR est caractérisée par les rubriques NUMÉRO-COMPTEUR, DIAMÈTRE-COMPTEUR, EMPLACEMENT, etc...

Dans la majorité des cas, chaque entité possède au moins une rubrique discriminante, c'est-à-dire une rubrique permettant de l'identifier sans ambiguïté.

Ex. : - La rubrique NUMÉRO-COMPTEUR est discriminante pour l'entité COMPTEUR
- La rubrique RÉFÉRENCE-ABONNÉ est discriminante pour l'entité ABONNÉ

L'entité est une sorte de « moule » général décrivant les caractéristiques que possède chaque occurrence particulière.

♦ Le concept de relation

Les informations du monde réel ne sont pas indépendantes. Leur dépendance se traduit par des relations.

Une relation représente une dépendance entre deux ou plusieurs rubriques. Pour définir une relation, le concepteur doit préciser la liste des rubriques concernées et le type de la relation. Il s'agit de déterminer les rubriques de départ et celle d'arrivée.

Ex. : les rubriques RÉFÉRENCE-ABONNÉ et DATE-RELÈVE (domaine de départ) sont en relation avec la rubrique INDEX (domaine d'arrivée).

Le type de relation caractérise la nature de la dépendance. Une relation est de type fonctionnel si à chaque valeur du domaine de départ correspond au plus une valeur du domaine d'arrivée.

Ex. : La relation entre RÉFÉRENCE-ABONNÉ, DATE-RELÈVE et INDEX est une relation fonctionnelle : pour un abonné donné et une date de relevé donnée, il n'y a qu'un seul index de relevé possible.

La relation entre NUMÉRO-COMPTEUR et INDEX n'est pas fonctionnelle : pour un numéro de compteur donné, il y a dans le temps, plusieurs valeurs d'index de relève. Une rubrique discriminante d'une entité est alors une relation fonctionnelle avec chacune des autres rubriques de cette entité.

Pour s'assurer de la fonctionnalité d'une relation, il convient de vérifier l'unicité de la valeur du domaine d'arrivée. D'où la notion d'identifiant.

L'identifiant d'une entité est un attribut, ou un ensemble d'attributs qui permet de repérer une occurrence d'une manière unique.

Autrement dit, parmi toutes les occurrences d'une entité, on ne peut pas en trouver deux qui aient la même valeur de l'identifiant : il ne peut y avoir deux **M. Alphonse NGABONZIMA, celui qui habite à tel endroit et qui est né tel jour** .

- Notion de cardinalités

Un aspect particulièrement important de la modélisation, expression d'une réalité, consiste à valoriser les liaisons entre entités et associations en termes de nombre d'occurrences minimum et maximum qui peuvent exister. C'est ce que l'on entend par cardinalités.

* La cardinalité minimum est le nombre minimum de fois où chaque occurrence d'une entité participe au type d'association.
* La cardinalité maximum est le nombre maximum de fois où chaque occurrence d'une entité peut participer au type d'association.

Ex.



. un enfant est fils d'une femme et d'une seule (1,1)
. une femme peut n'avoir d'enfant (0 enfant) ou en avoir plusieurs (n enfants)

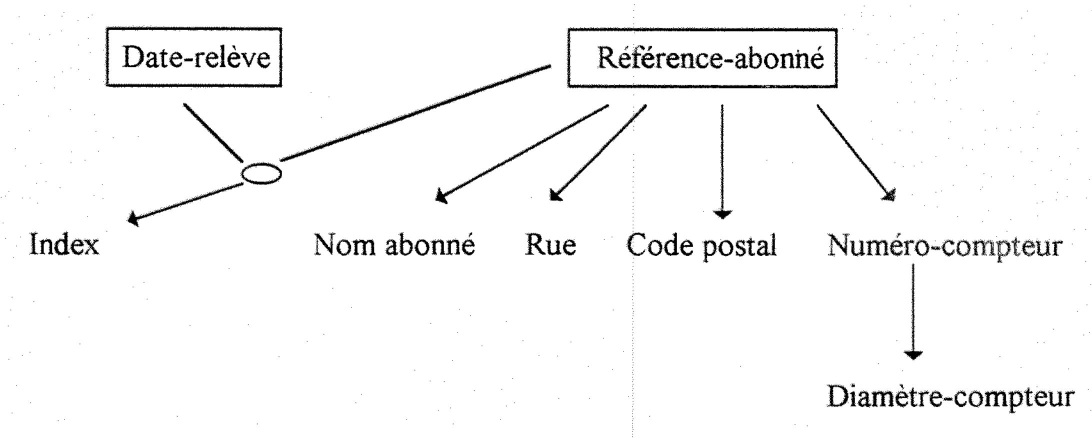
- Le concept de structure d'accès

Le concepteur doit pouvoir définir la façon dont sont organisées les rubriques du monde réel à l'intérieur du système d'information. Il lui faut un concept permettant de représenter la structure de l'ensemble de ces rubriques, c'est-à-dire des chemins autorisant l'accès à une rubrique à partir d'une autre.

Pour un système d'information donné, plusieurs structures d'accès possibles se différencient par les chemins d'accès.

Ex.: Prenons les rubriques : Référence-abonné, nom-abonné, rue-code postal, numéro-compteur, date-relève, index.

En tenant compte des relations fonctionnelles existant entre eux, l'une des structures d'accès possible est la suivante :



Cette structure montre qu'il est possible par exemple : d'accéder à la rubrique index à partir des rubriques Référence-abonné et Date-relève.

Le concepteur doit préciser les rubriques qui sont point de départ des chemins d'accès.

2. Le modèle de traitements

Alors que le modèle de données permet de décrire les données et leurs liaisons, le modèle de traitement doit permettre au concepteur de décrire le fonctionnement de l'organisation, à savoir les traitements effectués sur les informations gérées dans un système d'information. Mais ce modèle n'indique pas qui doit faire ni quand il faut faire.

Deux concepts : l'activité et le fonctionnement permettent de représenter ce qu'il faut faire fonctionner et sous quelles contraintes.

- Le concept d'activité

Une activité est un ensemble de traitements automatisés ou non, concourant à la réalisation d'une même finalité.

Ex. : L'abonnement est une activité d'un système d'information d'un service des télécom par exemple. Mais cette activité peut se décomposer en d'autres activités comme : abonnement-nouveau, reprise abonnement, mutation.

Le catalogue des activités du système d'information se présentera donc sous une forme arborescente dont le niveau de décomposition sera laissé au libre arbitre du concepteur.

- Le concept de norme de fonctionnement

Le concept de norme de fonctionnement permet de décrire des contraintes telles que :

- la durée d'exécution des fonctions et leurs conditions d'activation,
- les règles d'enchaînement des fonctions, c'est-à-dire la précédence, la simultanéité, l'implication ou l'interdiction entre fonctions,
- l'utilisation des ressources en qualité et en coût,
- la responsabilité des processeurs, c'est-à-dire des personnes ou les services, vis-à-vis des fonctions.

Cinq concepts permettent de décrire le fonctionnement du système sous forme de graphe de circulation :

- L'opération : c'est un traitement élémentaire, c'est-à-dire non décomposable qui a souvent une répercussion sur le contenu de la

structure elle-même. Elle correspond au plus petit niveau de décomposition de la fonction.
- Le lot : c'est un ensemble de rubriques apparaissant sur un même support en entrée ou en sortie d'une opération.
Dans une opération, un lot d'entrée contient les informations nécessaires à son exécution, un lot de sortie constitue un résultat de l'opération.

Ex.: L'opération **contrôle d'une commande** utilise trois lots d'entrée : bon de commande, produit, client et possède quatre lots de sortie : produit, bon de commande accepté, bon de commande en attente de stock et bon de livraison.

Le nombre d'occurrences d'un lot indique le mode de traitement de l'opération.

- Le lieu : c'est un ensemble de lieux physiques dans lesquels des opérations identiques sont exécutées. Il faut pouvoir préciser l'endroit où s'effectue chaque opération.

Ex. : le magasin, le service facturation, le service informatique, la réception sont des lieux.

- Le processeur : c'est la personne, l'équipe ou le service qui assume la responsabilité de l'exécution d'une ou plusieurs opérations.

Ex. : le **magasinier** est responsable de l'opération mise à jour des quantités de produits disponibles en stock. Le service **facturation** est responsable des opérations concernant la fonction de facturation.
- La ressource : est un constituant du système d'information dont une opération a besoin pour être effectuée.

Ex. : Un micro-ordinateur, une imprimante, le papier préimprimé des factures sont des ressources.

B - <u>Le dossier standard</u>

Le dossier standard est constitué de l'ensemble de paramètres caractéristiques de tous les concepts du module. On se souviendra que le rôle du concepteur est d'assurer pour chaque réalisation de concept une valeur à chacun des paramètres caractéristiques du concept.

Ex. : le concept de lieu est caractérisé par trois paramètres : le code, la désignation et le nombre d'occurrences.

Le concepteur devra, chaque fois qu'il définit un lieu, préciser la valeur du code, de la désignation et du nombre d'occurrences.

Le dossier standard est donc un guide permanent pour le concepteur, puisqu'il contient tout ce que doit définir le concepteur pour caractériser avec précision une réalisation du concept.

C - Les méthodes

En plus du modèle de données, de celui de traitements et d'un dossier standard, il faut que le concepteur ait des règles lui permettant de choisir, pour chaque système d'information, la chronologie d'étude des différents concepts.

D - Les fonctions automatisées

Lorsque le système d'information à concevoir est de taille importante, il devient très fastidieux de gérer manuellement l'ensemble des catalogues et de s'assurer de leur cohérence (élimination des redondances, des synonymes, des polysémies, etc...)

D'où nécessité d'un logiciel de conception assistée notamment pour :

- guider le concepteur lors de la saisir des valeurs des paramètres
- contrôler la cohérence des catalogues
- contrôler le bon déroulement de la méthode retenue.

Un tel logiciel permet d'obtenir un dossier de meilleures qualité et un gain de temps considérable.

NOTE :

L'objet de cette synthèse n'était pas de faire -loin s'en fallait - une étude exhaustive sur la conception d'un système d'information. Le domaine est suffisamment éloigné de notre compétence.

Il était tout simplement question de présenter très globalement le contenu de quelques documents qui ont traité de cette matière.
Une façon d'inviter notre éventuel lecteur à excuser les lacunes, les imperfections et les omissions dans cette synthèse.

III - MÉTHODOLOGIE DE LA RECHERCHE

III.1 - Recherche préliminaire

Afin de pouvoir cerner notre sujet dont la compréhension ne nous était pas très évidente, nous avons démarré notre recherche par la consultation des documents ci-après de la bibliothèque de l'ENSSIB :

- Encyclopeadia Universalis Corpus 19 : production - Rhumnales.- Paris : Encyclopeadia Universalis Éditeur, 1990 - ISBN 2 - 85229 - 287 - 4.

- Grand dictionnaire encyclopédique Larousse.- Vol. 9 : Relais à synchronie.- Paris : Librairie Larousse, 1985 - ISBN 2 - 03 - 102309 - 8

Les définitions et explications offertes par ces documents, si intéressantes qu'elles étaient, nous déroutaient un peu. Le concept de représentation des connaissances notamment, embrassait des domaines de la philosophie, de la linguistique, tellement que le lien avec la conception d'un système d'information n'était pas net.

Il nous a fallu le support de cours sur la conception d'un système d'information, document fourni par M. Mohamed HASSOUN, pour mieux circonscrire le champ de notre recherche et cibler les mots-clés.

III.2 - Détermination des mots-clés

Sur proposition de mon directeur de recherche, il a été décidé que les mots-clés qui guideraient notre recherche seraient principalement :

- Représentation des connaissances
- Structures de données (informatique)
- Data structure
- Base donnée
- Structuration
- Système innformation

Évidemment, cette liste de mots-clés a été augmentée par ajout des descripteurs ci-après trouvés sur les notices jugées pertinentes pendant l'interrogation des CD-ROM Pascal 1994.- Vol 1-2 et Pascal 1995.- Vol. 1-2 d'abord, et d'autres bases de données par la suite :

- Conception système
- Modélisation
- Base connaissance
- Banque donnée, etc...

III.3 - Recherche manuelle

Conscient que la bibliothèque de l'ENSSIB contient un fonds documentaire non négligeable en matière de sciences de l'information - pôle d'excellence de la BNF dans ce domaine - c'est principalement dans cet établissement que nous avons effectué la recherche.

En vue de recevoir, ne fut ce que quelques titres d'ouvrages qui nous mettraient sur les rails, la recherche manuelle s'est entamée par le feuilletage des documents ci-après :

- Les livres disponibles 1995 : french books in prints, sujets.- Paris : Éditions du Cercle de la Librairie, 1994.- Vol. 1-2.
ISBN 2 - 7654 - 0574 - 3

- Livres de France.- Paris : Éditions du Cercle de la Librairie, 1995.- Mensuel
ISSN 0294 - 0019

- Livres de France.- Paris : Éditions du Cercle de la Librairie, 1996.- Mensuel
ISSN 0294 - 0027

- Bibliographie National Française : livres.- Paris : Bibliographie Nationale de France, 1995.- Mensuel
ISSN 1142 - 3250

Les livres disponibles livrent des notices par thèmes ou sujets suivant l'ordre de la classification décimale de Deuvey, avec un classement alphabétique par nom d'auteurs.

Il était alors question de faire un parcours séquentiel des auteurs dans le thème « sciences de l'information » pour repérer des titres proches de notre sujet.

Nous avons pu y découvrir quelques ouvrages dont certains sont disponibles à la bibliothèque de l'ENSSIB. Disons toutefois que le fait que le thème soit trop global entraîne une lenteur dans la recherche. L'ordre alphabétique des auteurs pèche justement contre la proximité des titres à sujets proches.

Dans les livres de France, sous la rubrique « livres du mois », sont répertoriés les nouveautés et les nouvelles éditions. De même que dans la bibliographie nationale française : livres, nous avons pu y découvrir quelques titres.

L'étape suivante a consisté à dépouiller les bibliographies de ces ouvrages disponibles à l'ENSSIB. Soulignons évidemment que nous avons continuellement fait recourt à la recherche manuelle pour vérifier les documents obtenus par d'autres moyens de recherche et dont nous doutions de la pertinence.

III.4 - Recherche automatisée

III.4.1 - Consultation des CD-ROM

Afin de sélectionner les CD-ROM à consulter, nous nous sommes référés à:
- la liste des CD-ROM consultables à la bibliothèque de l'ENSSIB. Sur cette liste, on retrouve en gros les domaines contenus sur chaque type de CD-ROM.
- Annuaire du CD-ROM : les titres français/Jeannick Scolary -5e éd.
- Paris : A jour, 1994.- Réalisation [du] Centre documentaire d'A JOUR ISBN 2 - 903685 - 56 - 8

Cet annuaire recense les CD-ROM par thème. Nous l'avons à la bibliothèque de l'ENSSIB.

Tous les CD-ROM qui nous intéressaient sont disponibles à l'ENSSIB.

Pour nous assurer que nous n'aurions pas laisser tomber des bibliothèques qui pouvaient nous être d'un complément utile, un passage a été fait à la bibliothèque universitaire de Lyon I et à la bibliothèque municipale de la Part Dieu. Elles n'ont pu offrir aucun apport nouveau.

- Le CD-ROM Pascal :

C'est une série de CD-ROM, 2 volumes mis à jour par an. Elle contient la base de données PASCAL du CNRS.

Base de données multidisciplinaires, elle couvre la littérature scientifique et technique dont les sciences de l'information.

Le nombre de notices dépasse 3 000 000 dont les titres de périodiques, thèses et congrès.

La collection de la bibliothèque de l'ENSSIB couvre la période de 1987 à 1995.
La recherche s'est faite principalement à partir des mots-clés tirés du libellé de notre sujet. Nous étions obligés de les utiliser parfois en unitermes, d'autres fois sous forme d'expressions utilisées pendant l'indexation PASCAL du document.

Le système est conçu tel qu'on ne doit pas employer des conjonctions, mais nous procédions aussi par la combinaison de type booléenne pour être sûr d'avoir épuisé toutes les possibilités.
Critères d'accès aux notices catalographiques :
- Titres + Résumé + mots-clés
- Auteurs
- Mots-clés
- Langues
- Types de documents
- Adresses
- Sources
- Code de classement

Le tableau ci-après montre quelques unes des opérations construites :

Critères	Équations	Résultat	Réf. pertin.	Réf. choisies
Titre + mots-clés	DXF = Structuration	23	5	0
	DEF = Représentation ET DEF = connaissances	106	8	2
	DEF = Structure ET DEF = Données	3	3	3
	DEF =Data ET DEF = base	10	1	1
	DEF = Modularité ET DEF = connaissances	0	0	0
	DEF = Stockage ET DEF = information	56	3	0
	(DEF = Système ET DEF = information) ET LI = conception	109	18	4
	(DEF = Stockage ET DEF = structuration) ET DEF = Information	0	0	0
	(DEF = Organisation ET DEF = traitement) ET DEF = information	12	3	1
	Total après tri			11

- <u>Doc-Thèses</u>

Réalisé par l'ABES, Agence Bibliographique de l'Enseignement Supérieur, « Doc-Thèses » permet d'effectuer des recherches sur plus de 270 000 références de thèses françaises et de les localiser dans les bibliothèques. Sa mise à jour est semestrielle.

<u>Critères d'accès</u>

- Auteur
- Mot du titre
- Mot-clé
- Discipline
- Etablissement
- Année de soutenance
- aire géographique
- Directeur de thèse

etc...

Au début, nous avons choisi d'interroger en passant par la discipline. Mais comme le cheminement était long et sans résultats apparents, nous avons carrément opté de passer par les mots-clés et de dérouler les longues listes de titres pour les visualiser.

Cette deuxième option nous a permis de récolter quelques références pertinentes.

Critères	Mot-clé	Résultats	Réf. pertin.	Réf. choisies
Mot-clé	- Système information - Représentation des connaissances - Conception système information	284 40 1	16 1 1	9 1 2
	Total après tri	_	_	12

CD-ROM Lisa plus

Ce CD-ROM donne accès aux deux plus importantes bases de données concernant les bibliothèques et la science de l'information depuis 1969 : plus de 141 400 notices pour « Lisa » (Library and Information Sciences Abstracts) et plus de 5 900 pour « current Research in Library and Information Science ».
Le langage d'interrogation est l'anglais.
Fielded search

* ft = free text
* su = subject description
av = Author/Research worker
kt = Title keyword
ti = Title keyword
cn = comparate name
ab = abstract
la = langage
da = publication date

ed = editor/publisher
an = abstract number
kw = keyword
sf = subfile indicator
af = author affiliation
cy = contry Research
th = Thesaurus term
* cs = combine set

Ces critères précédés d'astirisque sont ceux utilisés dans la recherche.

Équation	Résultat	Réf. pertinents	Réf. choisies
1. (ft = design and ft = system) and ft = information	1721	_	_
2. (ft = Représentation and ft = Knowledge) and ft = information	291	–	–
3. cs = 1 and 2	41	5	0
4. (su = design and su = system) and su = information	5	2	1
5. (su = représentation and su = Knowledge) and su = information	2	1	1
6. su = data structure	5	–	–
7. su = data base	675	_	_
cs = 6 and 7	0	0	0

Total après tri	-	-	2

Nous avons visionné les 52 références et 8 seulement nous ont paru vraiment pertinentes. Il est possible que la non-maîtrise de la langue anglaise peut nous avoir entraîné à en éliminer d'autres importantes. Les textes sont mal évalués quand on ne maîtrise pas la langue ; de surcroît quand le sujet n'et pas de la spécialité du « documentaliste ».

- Le CD-ROM Electre-Biblio

Edité par le Cercle de la Librairie, Electre est une bibliographie de plus de 330 000 titres disponibles publiés en langue française, avec 12 000 références d'éditeurs et diffuseurs de la langue française. La mise à jour est mensuelle.

Critères d'accès aux notices

- Mots du titre et du sous-titre
- Descripteurs de vedettes-matières
- Mots du résumé
- Index général
- Mots-clés de Deuvey
- Table Deuvey
- Auteur
- Collection
- Éditeur
- Date
- Prix

Nous avons choisi une interrogation pour descripteurs de vedettes-matières.

Critères	Équation	Résultats	Réf. pertinentes	Réf. choisies
Descripteur de vedettes-matières	1. Représentation ET connaissances	0	-	-
	2. Représentation	9	-	-
	3. Connaissances	195	-	-
	4. (Système ET information) ET conception	13	2	1
	5. (Structure ET base) ET données	0	-	-
	6. Modélisation ET donnée	0	-	-
	Total après tri	-	-	1

- Le CD-ROM BNF (Bibliographie Nationale Française)

. Réalisé à partir de la base bibliographique BN-Opale de la Bibliothèque Nationale, il offre un catalogue de plus de 450 000 ouvrages entrés par le dépôt légal dans la partie officielle de la bibliographie de la France depuis 1975.

Les critères d'accès aux notices

ti = titres	*au = auteur
ap = auteur personne physique	ac = auteur collectivité
at = auteur-titre	ed = éditeur/imprimeur
li = lieu de publication	in = ISBN
co = collection	mc = mot notice
*mt = mot titre	ma = mot auteur
*ms = mot sujet	me = mot éditeur/imprimeur
mo = mot collection	an = date de publication
la = langue	cs = combiner les recherches

Les critères précédés d'astérisque sont ceux utilisés dans la recherche.

le CD-ROM BNF a été consulté après les autres. Il n'a pu nous fournir que des titres d'ouvrages que nous avions obtenus ailleurs. Aucun apport nouveau.

III 4.2 L'interrogation en ligne des bases de données

• Le choix des bases

La sélection des bases à interroger s'est faite à partir de la consultation des répertoires ci-après :

-Répertoire des banques de données professionnelles : banques et services d'information en ligne 1963 -1993. - Paris : ADBS Editions, 1993 . - 440 p
ISBN 2-901046-56-8

- On line bibliographic data bases : a directory and source book / James L.Hall .- 4e.- éd.- London : ASLIB, 1986.- 600p
ISBN 0- 85142 - 202 - 2

En plus de l'importance de l'information que la base était sensée contenir, le choix était aussi fonction de l'accessibilité.

•<u>La base pascal</u>

Cette base couvre la littérature internationale dans tous les secteurs de la science,dont les sciences de l'information, de la technologie et de la médecine.

Elle est une production de l'Institut National de l'Information Scientifique et technique (INIST).
Le volume de la base est de plus de 8.000.000 de références avec une augmentation de 4.5000 références par an.

En interrogeant cette base, nous envisagions compléter l 'information obtenue avec le CD-ROM.
Hélas nous n 'avons pu obtenir que des références que nous avions déjà eues par le biais du CD-ROM.

Accessibilité : Serveur Dialog, Questel

•<u>La base Loris</u>

La base Loris est chargée sur un gros ordinateur IBM de l'ENSSIB. Elle contient plus de 8000 références bibliographiques de monographies,thèses, mémoires,rapports et numéro d'articles. Son logiciel (Doris) permet le déchargement des notices.

La bibliothèque de l' ENSSIB ayant un fonds spécialement tourné vers les Sciences de l'information ,cette base nous a fourni 15 notices pertinentes sur 62 capturés.
Il est précisé que les termes de recherche sont les termes de la liste RAMEAU.
Les mots-clés utilisés pour la recherche sont ceux qui nous avaient donnés de bons résultats lors de la consultation des CD-ROM..Voici certaines des réponses obtenues:

9 Représentation des connaissances
38 Système information
2 Base connaissance
35 Base données
12 Structures de données (informatique)
19 Modélisation
1 Structure information

• <u>La base Sibil</u>

Cette base a été développée en 1971 par la bibliothèque cantonale et universitaire de la bibliothèque cantonale et universitaire de Lausanne. Le réseau SIBIL(Système Intégré pour les bibliothèques universitaires de Lausanne) regroupe aujourd'hui 32 bibliothèques interuniversitaires et universitaires et des bibliothèques d'organismes de recherche français.
Elle contient plus de 800.000 notices. Plusieurs types de documents y sont représentés à savoir: les monographies, les thèses, les mémoires, le congrès, la littérature grise etc .

C'est aussi le cas de « Structuration »dans la base pascal. On pouvait avoir des titres tels que:

ex : Structuration cognitive et logique.

Pour recevoir des résultats probants, il était préférable de passer par l'interrogation de type booléen.

Mentionnons également que l'interrogation sur le champ résumé est lui aussi source d'un fort taux de bruits.

Analyse du Silence

Même si l'idéal serait d'être le plus possible exhaustif, il n'est pas facile de le réaliser. Il existe toujours des paramètres limitatifs. Pour l'interrogation en ligne notamment, on doit tenir compte du coût d'utilisation. En raison de quoi nous étions obligés , pour l'interrogation de la base pascal par exemple, de n'utiliser que des descripteurs qui avaient de bons résultats quand nous consultions les CD-ROM.Ceci nous amène à supposer qu'il y'a des descripteurs qui auraient pu offrir d'autres bons résultats mais que nous n'avons pas osé exploiter. Etant donné que nous n'étions pas parvenus à avoir les thésaurus de ces bases,il n'était pas non plus évident de pouvoir formuler correctement toutes les équations de recherche.

Localisation et obtention des documents primaires

Comme déjà dit ci-haut, la bibliothèque de l'ENSSIB dispose d'un fonds important en science de l'information. Beaucoup d'ouvrages notamment y sont disponibles.

Quant aux articles recensés dans pascal,ils sont localisés à l'INIST. Le service de prêt interuniversitaire de l'ENSSIB nous a aidé pour faire une commande d'articles (photocopies) jugés intéressants. Ici aussi, la commande étant payante, on ne pouvait se permettre que l'acquisition du strict nécessaire.

Le fichier catalographique est alimenté quotidiennement après validation des notices créées ou localisées par chaque bibliothèque du réseau.

.Interrogation de Sibil

Une fois le critère de recherche formulé en vue de la consultation du catalogue (catalogue principal) par matière, le système propose de choisir entre :

-Notice guide
-Ouvrages généraux
-Thèses

La recherche se fait par année de publication .Nous avons exploité les descripteurs qui suivent :

-Représentation des connaissances
-Structure de données(Informatique)
-Conception Système
-Système information

Sur 54 références capturés, 5 seulement ont été jugées pertinentes.

Synthèse des résultats

Banque de données	Resultats	Réferences pertinentes	% de bruits
Pascal	271	41	84%
Loris	62	15	75%
Sibil	56	5	91%

Analyse du bruit

Les mots-clés "représentation", "structure'' ,``structuration'',utilisés en uniterme, donnent lieu à un taux élevé de bruits. Ils changent de sens en fonction du contexte.

C'est pour cela qu' on pouvait avoir des résultats comme « structure sociale » qui n'ont rien à voir avec la conception d'un système d'information.

Dans les bases Sibil et Loris, « représentation des connaissances » donne lui aussi des résultats qui se rapportent plus à l'intelligence artificielle et à la linguistique qu'à notre sujet.
Voici quelques exemples pour illustrer notre propos.Le premier exemple est pris dans les notices capturées sur Loris , le 2e dans celles récupérées sur Sibil.

Ex.1 : L'Intelligence artificielle et le langage. Vol 1. Représenter des connaissances/Gérard Sabah. - Paris : Hermès, 1988.- 352p. : ill. ; 24cm.

Ex. 2 : Le contrôle dans les systèmes à base de connaissances. Contribution à l'épistémologie de l'intelligence artificielle / Bruno Bachimont. - Paris : Hermès, 1994.

BIBLIOGRAPHIE

Comme la plupart des documents que nous avons pu recenser traitent simultanément de plusieurs thèmes de notre sujet, nous avons jugé qu'il serait inexact d'essayer de faire un groupement thématique de nos références. Nous optons plutôt pour les regrouper par type de documents.

I. Ouvrages généraux

- BEYNO-DAVIES, Paul
Systèmes d'information :conception et mise en oeuvre .- Paris : afnor , 1992 .- XII - 232p .- ISBN 2-12-480811-7

- BOUSSARD, Jean Claude , MAHL , Robert
Programmation avancée : algorithmique et structures de données. - 4è éd. Paris : Eyrolles, 1987 . - 248 p. -
ISBN 2-212-08544-3

- BRAND, Fortner
The data handbook .- Second ed.- New-York : Springen Verlag, 1995.- XXII - 350p

- COLLONGUES, A., HUGUES, J., LA ROCHE, B.
Merise : Méthode de conception .- Paris : Dunod Informatique, 1990.

- CAVALERO, J.L., HERIN-AIME, D
La conception des systèmes d'information : un modèle, un dossier standard, des méthodes.- Paris:Masson, 1982.- ISBN 2-225-78689-9

- COURBON, Jean-Claude
Systèmes d'information : Structuration, modélisation et communication .- Paris : Inter- Editions, 1993.-286p

- COURTIN, Jacques
Initiation à l'algorithmique et aux structures de données .- Paris : Dunod, 1987.

-DELOBEL, Claude
Bases de données : des Systèmes relationnels aux Systèmes à objets .-1ère éd.- Paris : Inter Editions, 1991.- XVII -460p.

- DUNAUD, N., et . al.
Réussir un projet de conception orienté objet.-Paris : Masson, 1995.-261p

- FLORY, André
Bases de données : conception et réalisation.- 2° éd. Paris : economica, 1987.- 164p.
Bibliogr. pp 155-157

- FOUCAUT, O., THIERRY,O., SMAILI,K .
Conception des Systèmes d'information et programmation événementielle: de l'étape organisationnelle à l'étape physique .- Paris: InterEditions, 1996.-240p.-
ISBN 2-7296-0559-2

- GALACSI
Comprendre les Systèmes d'information: exercices corrigés d'analyse et de conception .- Paris: Dunod, 1985.-145p
ISBN 2-04-016480-8

- GALACSI
Conception de bases de données: du schéma conceptuel aux schémas physiques.- Paris : Dunod, 1989.- 378p
ISBN 2-04-018720-0

- GALACSI
Les Systèmes d'information: Analyse et conception .- Paris: Dunod, 1990.- XIV-304p.- ISBN 2-04-019607-2

- GUYOT,J. ,VIAL, C.
Arbres,tables et algorithmes.-Paris:Eyrolles, 1992.

- LAMPRON, R., et al.
Introduction aux Structures de données et aux fichiers.- Montreal, 1987.

- LE MEN, Yvan
une énergie nouvelle pour mieux communiquer Fichiers et bases de données.-Paris: Dunod, 1991.

- LEMOIGNE, Jean-Louis
La modélisation des systèmes complexes.-Paris: Dunod, 1990.-192p
ISBN 2-04-019704-4

- MIRANDA, Serge
Introduction aux bases de données.- Paris: Eyrolles, 1984.- XX-227p.

- MIRANDA, Serge
Les bases de données relationnelles. - Paris : Eyrolles, 1986.- XXIII - 354p.

- NIDAL, A., LALANNE, JC.
Architecture technologique des Systèmes d'information .- Paris: Les éditions d'organisation, 1989.

- PEAUCELLE, Jean-Louis
Les Systèmes d'information: la représentation en mémoire.- 2° éd.- Paris: INRIA, 1979. - 275 p. - ISBN 2-72610-229-8

- PLANCHE, Remi
Maîtriser la modélisation conceptuelle.- Paris : Masson, 1988. - 256p. -
ISBN 2-225-81203-9

- REYNIER, M.
Assurance qualité en conception: simplicité, bon sens et rigueur pour améliorer l'efficacité.- Paris: Masson, 1992.- 168p.-
ISBN 2-225-83965-4

- RUMBAUGH, J. et al.
Object Oriented modeling and Design.- Prentice . Hall, 1991.

- ROWLEY, J.E
The basics of Systems analysis and design for information managers.- C.Binglay, 1990.-158p

- SENN, James
Analyse et conception de Systèmes d'information / trad. de l'américain Raymond Borraz.- Mc Graw-Hill Editeurs, 1986.- 648p.-
ISBN 0-07-548774-8.

- TARDIEU, H., ROCHFELD, A., COLLETTI, R., PANET, G.
La méthode Merise, tome 2, Démarches et pratiques.- Paris: Les Editions d'organisation, 1985.

- TARDIEU, H., ROCHFELD, A., COLLETTI, R.
La méthode Mérise, Tomes: Principes et outils.- Paris: les Editions d'organisation, 1989.

- TARDIEU, H., NANCI, D., PASCOT, D. ,
Conception d'un système d'information : construction de la base de données. - éd. revue et corrigée. - Paris : les Editions d'organisation, 1984
ISBN 2-7081-05-96-5

II.THÉSES ET MÉMOIRES

- ALIMAZIGHI, B.
Conception et spécification de Systèmes d'information : analyse comparative de méthodes et utilisation du langage ADA. Th. Doct : Paris 6, 1986

- ALVARES, O.
Contribution à l'étude du pilotage de la modélisation des systèmes d'information. Doctorat d'Etat : Grenoble 1, 1988.

- AYACHE, M
Cassiopée : Vers une méthodologie de conception par réutilisation. Th. doct : INSA Lyon, 1984. - 238 p

- BARI, M.
Une méthode d'analyse et de conception orientée objet de systèmes d'information actifs. Th. doct : Paris 6, 1992

- BOUZEGHOUB, M
SECSI : un système expert en conception de systèmes d'informations, modélisation conceptuelle de schémas de bases de données. TH. doct. : Paris 6, 1986.

- BRUNE, J.
Analyse conceptuelle orientée objet. Th. doct. : Paris 6, 1993.

- CHRISMENT, C.
Conception d'un système d'informations automatisé : définition interactive d'un schéma de base de données. Th. doct. : Toulouse 3, 1980.

- CAUVET, C.
Un modèle et un outil d'aide à la conception des systèmes d'information. Doctorat d'Etat : Paris 6, 1988.

- DELOBEL, C.
Contribution théorique à la conception et à l'évaluation d'un système d'information appliqué à la gestion.- Grenoble : Université Scientifique et médicale de Grenoble, 1973.

-DURAFFOUR, F.
Analyse critique des systèmes d'information de gestion et de leurs principales approches méthodologiques. Th. doct. : Paris 9, 1993.

- GREVOBAL, M.-C.
La production d'explications, vue comme une tâche de conception : contribution au projet d'aide. Th. doct. : Université de Conpiègne, 1994. - 196 p.

- HUEY, B.
Le sujet, l'objet, le projet dans la modélisation des systèmes d'information. Th. doct. : Paris 6, 1992.

- MORAND, B.
Processus de conception des systèmes d'information avec un modèle d'acteurs : Essai sur le système de représentation. Th. doct. : Caen, 1994.

- PETER, P.
Méthodes de classification hiérarchique et problèmes de structuration et de recherche d'information assistées par ordinateur. Th. doct. : Rennes1, 1987.

III. ARTICLES ET COLLOQUES

- BONJOUR, H., et al.
Bases de concepts et intégration de bases de données. *Ingénierie des systèmes d'information*, 1995, vol 3, 10 2-3, pp 421 - 440. - ISSN 1247-0317

- COLLOQUE DE NAMUR
Data structures models for information systems ; procedings, may 27-30, 1974.- presses universitaires Namur, 1974. - 200 p. -

- ESPINASSE, B.
Merise + : une extension de Merise à l'approche objet par un apport de Hood. Ingénierie des systèmes d'information, 1995, vol.3, n° 2 - 3, pp 301 - 325 . -
ISSN 1247-0317

- GEFFROY, J.-C.
Identification des systèmes séquentiels : une approche unifiée. TSI. *Technique et Science informatiques*, 1995, vol. 14, n° 7, pp. 809-837.-
ISSN 0752-4072

- HABRIAS, H.
Les spécifications formelles pour les systèmes d'information quoi ? pourquoi ? comment ? *Ingénierie et systèmes d'information*, 1995, vol. 3, n° 2 - 3, pp 205-253.
ISSN 1247-0317

- HANNABUSS, C. S.
Knowledge representation and information exchange. Robert Gordon University School of Librarianship and information studies, October 1990.

- HYENNE , J.
Points clés pour l'intégration de données dans l'entreprise. IDT. *Information, documentation transfert des connaissances*, Congrès 12, Paris 13 - 15 Juin 1995, pp184-188. - ISSN1140-6992

- KLEIN, G.H
Multifactorial descriptive modelling on irregular data designs. *Mathematical and computer modelling*, 1995, vol. 22, n° 1, pp. 1 - 9
Bibliogr. 35 réf.
Résumé en anglais

- KOUM, G.
Vers des Schémas conceptuels de base de données issus d'un méta-modèle. *Rapport*,CNRS.Centre de Recherche informatique de Nancy, 1993. 2pp.

- LOMPRE, N
Analyse des requêtes dans le cadre de la conception d'un système à base de connaissances. Travail, 1994, vol 57, n° 1, pp71 - 88. -

- RAUH. O.
Relationship modes to cover modelling of deductive information system. *Informatik, 1995*, vol 10, n° 3, pp. 128 - 138.

- VATONNE, D., CHRISTINE, M.
Les termes : un modèle algébrique de représentation et de structuration de données symboliques. *Mathématique et sciences humaines*, 1993, n° 122, pp. 41 - 63. -

GLOSSAIRE

CD-ROM : Compact Disc Read Only Memory

BNF : Bibliothèque Nationale de France

LISA : Laibrary and Information Sciences Abstracts

SIBIL : Système Intégré pour les Bibbliothèques universitaires de Lausane

INIST : Institut national de l'Information Scientifique et Technique

ENSSIB : Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques.

ABES : Agence Bibliographique de l'Enseignement Supérieur

CNRS : Centre National de Recherche Scientifique

OPAC : Online Public Access Catalog (Catalogue Informatisée des Bibliothèques).

